Communication Technologies branch
OFFICE OF COMMUNICATIONS
NATIONAL CANCER INSTITUTE

6116 Executive Boulevard | Suite 3048A | Rockville, MD 20852
301.594.3330 | fx: 301.435.6069
http://usability.gov

Searching vs. Linking on the Web: A Summary of the Research

Sanjay J. Koyani and Robert W. Bailey, Ph.D.

July 19, 2002

Communication Technologies Branch

Office of Communications

National Cancer Institute

INTRODUCTION

This report summarizes the available research on how users search for information within Web sites. It addresses topics such as users' search behaviors, how to improve users' accuracy and success in finding information by using a site's search capability and site links, and when to encourage users to use site links versus the search capability to find information.

Web site users essentially have three ways to find content. They can read or scan a page using scrolling to get to new information, they can click on links and follow them to the desired content, or they can use the site's search facility. It is the latter two, linking and searching, that are addressed in this article.

Linking means that users will select and click on a link on a starting page (usually the homepage), which then causes a new page to load. Users continue toward their goal by finding and clicking on subsequent links (see Figure 1). This behavior continues until the desired content (i.e., a target item) is found.

Figure 1. Example of 'linking' to find information.

Searching means that users will access the search capability that has been provided on a web page. Generally, the search facility is shown in an entry field called a "search box" (see Figure 2). Users typically enter one or more words into the search box, which hopefully will match words used in the site. When there are words in the Web site that match the words submitted by users, the search results are shown.

Figure 2. Example of 'searching' to find information.

In general, when users are presented with a new page, they first tend to scan for words related to achieving their goal. Some of these words will be part of the content of the page, while others will be links to other places (usually within the Web site). If users do not find what they are looking for, then they may use the search capability. That is, when users cannot efficiently find appropriate links, they make up their own links by entering search terms into a search box.

This article addresses numerous issues related to linking and searching within a Web site.

RESEARCH FINDINGS

♦ More than Half of Users Appear to be Search-Dominant

Nielsen (1997) reported the results of some usability studies where he found that:

- More than half of all users are search-dominant,
- About 20% of all users were link-dominant, and
- The remaining 30% exhibited mixed searching and linking behaviors.

The *search-dominant users* were not interested in exploring a site, but went straight to the search facility. They seemed to be task-focused, and tried to find specific information as fast as possible. The *link-dominant users* used the links to move around a site -- even when they want to find specific information. These users tended to use the search facility only when they were hopelessly lost. The mixed-behavior users did not have a strong preference, and switched between using the search facility and clicking links depending on what seemed to be the best way to navigate a site.

♦ Preference for Search vs. Link Behavior Depends on the Web Site and Tasks at Hand

Spool and his colleagues concluded that it is not the predisposition of users, but more likely the type of Web site and the tasks being performed by users that determines whether or not most users search or use the links (UIE, 2001d).

Something inherent in the site's design seemed to cause users to choose the search facility or to select a series of links. In one study (UIE, 2001d), researchers found that:

- 21% of the sites elicited primarily search-dominant behaviors from users, where they mostly used the search facility,
- 32% of the sites elicited primarily link-dominant behaviors from users, where they mostly clicked on links, and
- 47% of the sites elicited mixed behaviors.

Their study had 30 users perform 121 different shopping tasks. Each user visited between 3 and 6 Web sites, and shopped for items that they were interested in purchasing. The researchers reported that *not one person* used the search engine first

when looking for product information, while about 20% chose links exclusively. Although many users in their study stated that they "always go to search immediately," none of them *did* use the search engine first.

In the study, nobody exclusively used the search facility on every site, while 18% *always* used links. On 21% of the tested sites, every test participant used the search facility at least once. This was most likely because of the way these sites were designed or the nature of the tasks that were being performed.

Based on the available research, it is not clear whether or not certain *users* are pre-disposed toward using the search facility or clicking on links. It seems more likely that the Web site itself, and the tasks being performed, are much stronger predictors of users selecting either to search or to use the links.

♦ How to Determine 'Link' Vs. Search Behavior

Certain Web site elements seem to influence whether users search or link. One element was the type of product being sold. For example, specific book or CD titles encourage searching, whereas looking for non-specific items, like clothing, seem to encourage linking.

Figure 3. An example of a specific Web topic that might invoke users to search first.

If users are looking for something specific, and they know the item number or precise title, they tend to use the search facility (Figure 3). When users are looking for an item within a category of items (e.g., a recent report on thyroid cancer, a blue pullover sweater), they attempt to use links to find the item of most interest (Figure 4).

Figure 4. An example of a non-specific Web topic that might encourage users to use links first.

Spool (2002) found that the decision to use the search facility depended on whether the user was looking for items that are totally unique (finding a specific book), or an item that is similar to many other items (finding a good digital camera).

For items with totally unique information, such as titles of CDs, videos, or books or product numbers, claim numbers, tax forms, etc., the search facility is used more and is more likely to succeed. For items that are not totally unique, such as computer accessories, apparel, corporate policies, news and health topics, the search facility tends to be used less, and is less likely to lead to a successful outcome.

A second element that seems to influence whether users search or link was having a Web site that encouraged the initial use of a search facility. For example, a site that primarily provides users with a search box and relatively few link options on the home page. It was found that some people initially searched to find links, and then used linking behavior until they found their goal.

A third element was related to unusable links. Some users began searching only after the links proved not to be too useful. In other words, the search facility was used as a fallback after users failed to pickup 'scent' on a homepage.

♦ 'Linking' More Effective than 'Searching' in Finding Content

After examining data from hundreds of users on dozens of sites, one set of researchers found that linking led to more successes than did searching (Ojakaar and Spool, 2001). Their study reported that users found their desired content far more often when they navigated by using links rather than using the search capabilities. In fact, users were 77% more likely to find their target when they used linking. In their study, users who used the search facility were successful only 30% of the time, while those who used links were successful 53% of the time.

Using links has the obvious problem of users not being able to understand what the terms used in the links mean. But users can run into even more problems when they try to find content by using a within-site search facility. For example, one user tried to find out about Amazon's "return policy." She found 43 books on the topic, but nothing on Amazon's return policy.

♦ Presenting More Links Increases the Likelihood That Users Will Understand Site Organization and Terminology

The more familiar users are with a site's terminology and organization, the more likely they will use links, rather than searching (Ojakaar and Spool, 2001). When most users have little understanding of a site's terminology or organization, designers should present more links, which means they are presenting many different 'trigger' words (see Figure

5). When user's have a high level of familiarity with a site's terminology and organization, designers can get by with fewer links.

Figure 5. An example which includes many links to help users navigate the Web site.

♦ Links Elicit Less 'Clicks' Than Search

Spool (2002) studied the click patterns of users who knew exactly what they wanted *and found it.* Users who used the search facility took an average of 5.1 clicks, while those who followed links took an average of 4.4 clicks. The people who searched made about 14% more clicks than those who only followed the links.

♦ Linking Leads Users to View More Related Content

Spool and his team explored another interesting search-related issue (UIE, 2001a). They attempted to determine whether 'Searching' or 'Linking' through to a target would best lead users to explore related content. After observing 30 people attempting to find items in Web sites, and then watching them look at other content, they noted that Linking, more than Searching, resulted in exploring related content.

If users searched to locate their target, only 20% continued looking at other content after they found their target. However, when users used links to locate their target, 62% continued to browse the site (see Figure 6). A closer inspection of their data showed that users were three times more likely to find related (valuable) content if they linked through to a target rather than using a search facility. In fact, if users started by using links on the homepage, they ended up looking at almost 10 times more non-target content pages than those who started by searching.

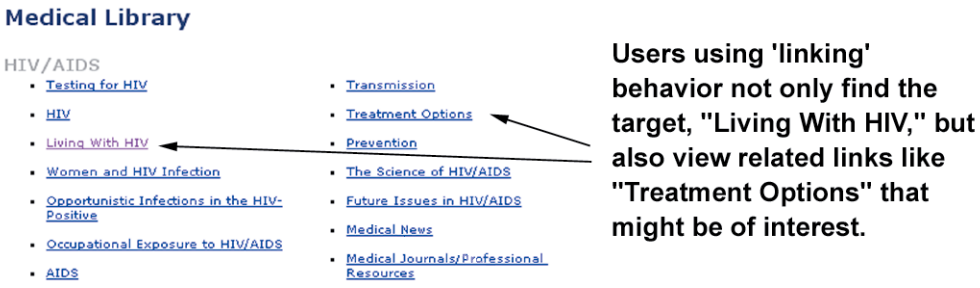

Figure 6. An example of how linking can help users see additional related information.

♦ If Users Do Not Have a Successful Search Experience on their First Attempt, They May Turn to an External Search Engine

Nielsen (2001) reported the results of a study where a large group of people was observed shopping on various e-commerce sites. Their search success rate on their first search was 51%. When they elected to change the keywords and search again, their success rate went down to 32%. On their third search, the success rate was only 18%. He concluded that if users did not find the result they needed with their first query, they are progressively less and less likely to succeed with additional searches. In his study, almost half the people whose first search failed gave up immediately.

More recent studies (UIE, 2001b; UIE, 2001c) also observed users trying to search various sites. Consistent with previous findings, they found that users found their target only 55% of the time on their first try. When users tried a second time with different keywords, their success rate dropped to 38%. In these studies, users *never* found their target when they conducted three or more searches using different keywords. The study results are summarized in Table 1.

	Successful Searches	
Search Attempts	**Nielsen**	**UIE**
First	51%	55%
Second	32%	38%
Third	18%	0

Table 1. A summary of the studies showing first, second and third attempts at searching.

Nielsen (2001) argues that typical users are poor at query reformulation. This means that they are not willing or may not be able to refine their queries. When they do not get good results on their first try, the second and third search attempts rarely succeed. Designers should do whatever they can to increase the users' success on their first search.

In the UIE studies, only 25% of the users searched more than two times. The main reason that people tried to search more than once was because they received the "no

results" message. Some designers provided search tips, such as "Try a new search using different terms," to help users who failed to find what they were looking for. Their studies suggested that providing these tips had no positive effect on their subsequent efforts.

If users do not have a successful search experience within a Web site, they may quit using the site's search facility in favor of an external search engine such as Google, AltaVista, Yahoo, Excite or HotBot. This, of course, means that they will be searching for the same information across many sites, which will give them the opportunity to go to another Web site.

♦ **Designers Can Improve Search Success by Planning for User Input Errors**

Spool (2002) has provided some insights as to why searches fail. He observed several people conducting searches on numerous Web sites, and found that:

- About one-third of the searches actually helped users find their desired information,
- Another one-third of the searches provided non-useful information to users, and
- One-third provided the message that "no results" were available.

Spool reported that about half of the unsuccessful searches were the result of users making input errors. These included misspellings (67%), using inappropriate case, spaces and punctuation (17%), misusing plurals (8%), and making typing errors (8%). These results are shown in Figure 7.

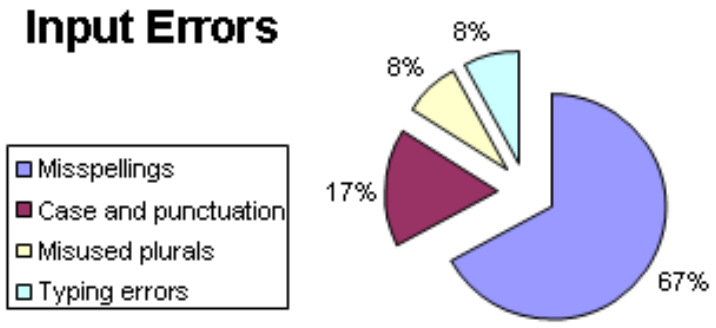

Figure 7. The relative proportion of input errors when entering search terms

In most of these cases, designers could have taken steps to ensure that users were more successful. Designers should have:

- Determined common misspellings and allowed them to be accepted by the computer,
- Allowed both singular and plural forms of commonly used keywords,
- Allowed both upper- and lowercase letters and common forms of punctuation to be accepted, and
- Determined the most common typographical errors and accepted them as correct.

A second major problem, one that resulted in about one out of five searches being unsuccessful, was the lack of a match between users' words and designers' words. In these situations, the user and the designer used different names for the same content. Again, designers should (a) determine the words that users will most likely use, and (b) allow those words to be acceptable as synonyms to the computer. All words (search terms) used by users should be collected, stored, and when possible, matched to the target being sought.

Spool (2002) reported that users' search terms were not present on the page 71% of the time. This means that on about 3 out of every 4 pages, users could not find the terms necessary to continue linking toward a goal. When the desired terms were not found, users are more likely to use the search capability. About half of the user entries into the search box were general terms, and not terms that would quickly lead them to the final goal.

♦ Search Templates Improve User Accuracy

One way to improve search performance is to supply sets of keywords for searchers. Fang and Salvendy (1999) developed a template to help users formulate a search. The intent of the template was to assist users in the selection of keywords used for searching. The feature was organized as a hierarchy of predefined keywords that would:

- Help to restrict the users initial search sets, and
- Improve the relevance of the "hits."

Several predefined search queries (i.e., sets of keywords) were created by subject matter experts and Web searching experts (see Figure 8).

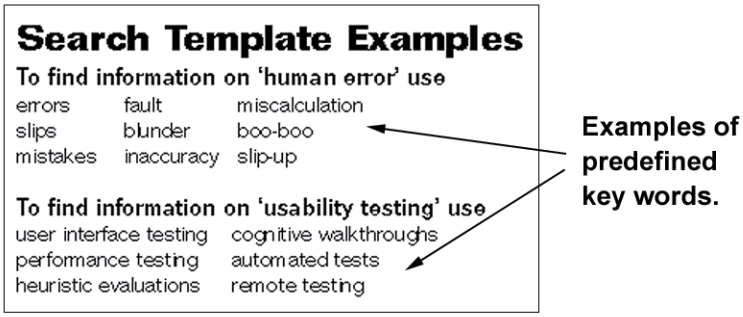

Figure 8. Examples of a search template.

They had 10 participants search for home pages that contained a predefined word. The predefined word could not be used as a keyword in a search query. The participants could use any search strategy to find as many relevant home pages as possible in a limited time period.

Participants using the search "templates" found 70% more Web sites ($p<.01$) than those using a traditional search engine. There was no reliable difference in user satisfaction between the two search engines. They concluded that the feature did facilitate the search process by helping users better formulate their search queries.

♦ To Improve Search Engine Hit Rate, Survey Users for Their Search Terms

As was mentioned earlier, another way to improve search performance is to have designers do a better job of creating the words that are used in Web sites. Users tend to use words and phrases that are different than those used by designers. Users obviously do not think exactly like designers do, and usually do not even think like each other.

It can be difficult for a user to guess the exact words used by designers of a Web site. Even so, when users are attempting to search a site for desired content, they generally must use the same words that were selected by designers of the site. What are the chances that users will guess the exact same word that designers elected to use? These probabilities were calculated in a study conducted a few years ago at Bell Laboratories (Furnas, 1984).

The following is an estimate of "hit rates," where users and designers agree, when users search with their keywords. Designers determine the Web site terms:

- Without conducting any studies: Up to 20%,
- By conducting a survey of potential users: 20% to 40%, and
- By conducting a survey of potential users, and allowing up to five synonyms for each term: 40% to 60%.

It is possible for designers to increase the hit rate up to 95% by conducting a survey of potential users, and then continuing to collect the words used by users as they use the search facility within a site. To accumulate a set of terms, the Web site must maintain a list of terms that users attempt, and the resulting action with which they are eventually satisfied (where the search is successful). The computer then can automatically respond to one or more search terms by providing users with a set of alternatives. Users then select the alternative that seems to best represent their needs.

♦ Consider Natural Language Search Capabilities to Improve User Performance, Preference

Natural language search engines allow users to type in complete questions instead of one or more keywords. For example, using a traditional search capability, users could type keywords such as "Jim Thorpe." This usually provides users with much more information than they want. Using a natural language search capability, users can ask specific questions in plain English, such as "What year did Jim Thorpe receive his first gold medal in the Olympics," and have a high probability of getting an exact answer.

Natural language search capabilities retrieve answers to specific questions asked by the user, which can be (a) a single response, (b) a list of possible responses, or (c) one or more questions to help narrow the search.

In addition to analyzing the meaning of words, grammar, and syntax, some natural language systems attempt to find patterns in the way questions are asked. Most natural language search engines attempt to learn from every question asked by users. As users participate in the question-and-answer process, these search engines attempt to adapt and improve.

Natural language search engines provide an unconstrained medium for users to freely express their wants, needs, and interests. Analyzing users' queries enables organizations to know more specifically what kinds of questions visitors are asking and what kind of information they are looking for.

This following study evaluated the JeevesOne search capability. Others offering similar search applications for organizations include AltaVista (Palo Alto, California) and Autonomy (Cambridge, England).

A comparison of natural language (JeevesOne) searching and traditional searching was conducted by the eTesting Labs in April 2002 (JeevesOne-2002). They had 209 people use both a traditional search capability and a natural language search capability to answer a series of questions. The traditional search capabilities already were on existing, live Web sites (prudential.com and bea.com). They modified JeevesOne, a natural language search capability so that it could adequately search the information in these two Web sites.

Each participant performed four searches, two using the traditional method and two using the natural language method. Each person was required to complete each search in five minutes or less, and was allowed to "give up" prior to the five-minute deadline if they could not find the answer. After completing their searches, the subjects responded to a questionnaire.

The results are shown in Table 2 and Table 3.

Search Method	Percent Finding the Answer	Percent Giving-Up	Average Time to Complete	Average Keystrokes and Clicks	Average Pages Required
Natural Language	91%	2%	136 sec.	26	3.8
Traditional	55%	16%	189 sec.	30	6.8

Table 2. Performance results from comparing natural language and traditional searching.

Search Method	Percent that Judged the Search Engine as Easy-to-Use	Percent Likely to Return to the Web site to Search	Percent that Judged the Product as Better than other Search Engines
Natural Language	75%	68%	63%
Traditional	38%	38%	27%

Table 3. Preference results from comparing natural language and traditional searching.

It appears that the natural language search capability used in this study allowed test participants to find the answers to their searches with more success and in less time. In addition, the participants were much less likely to give up searching. When compared with a traditional search capability, the subjects judged the natural language method as easier to use. In addition, they indicated that they were more likely to return to the Web site and use the search capability again.

Unfortunately, their study did not report whether or not any of these differences between natural language and traditional search capabilities were reliably different. Under these circumstances, we only can guess where true differences (if any) exist.

Currently, the cost of natural language search capabilities can be quite high. For example, one organization (a state government) now pays about $18,000 per month for a hosted application natural language search service. That is considerably more than the state was paying for using a keyword search technology. Natural language software licenses can range from $125,000 to $1.5 million.

♦ Certain Issues Have Little Impact on Search Performance

Spool (2002) suggested that user success is not affected by:

- Relevancy scores (users seemed to expect that all results are relevant),
- Sorting (users tended to ignore this functionality),
- Using advanced search capabilities, and
- Providing "search tips" (users rarely read the tips, and even more rarely applied the ideas).

Even though providing search tips may not work, the "search tips" that are usually suggested to users include:

- Check your spelling,
- Try using similar terms,
- Use fewer words,
- Avoid plurals (use mouse, not mice), and
- Use partial words instead of whole words (use record, not recordable).

♦ Consider Limiting Search Engines to Subsites as Web Size Grows

Nielsen (2001) observed that as Web sites continue to grow and offer multiple services in a single site, designers may want to consider offering search capabilities that are limited to one or more subsites. These subsites should be clearly identifiable to typical users (see Figure 8).

The default search scope always should be set to include the entire site. If a limited search returns too few or too many results, give users an easy way to expand or limit the search to other subsites. When users choose to limit the scope of a search, the limitations should be clearly indicated at the top of the results page.

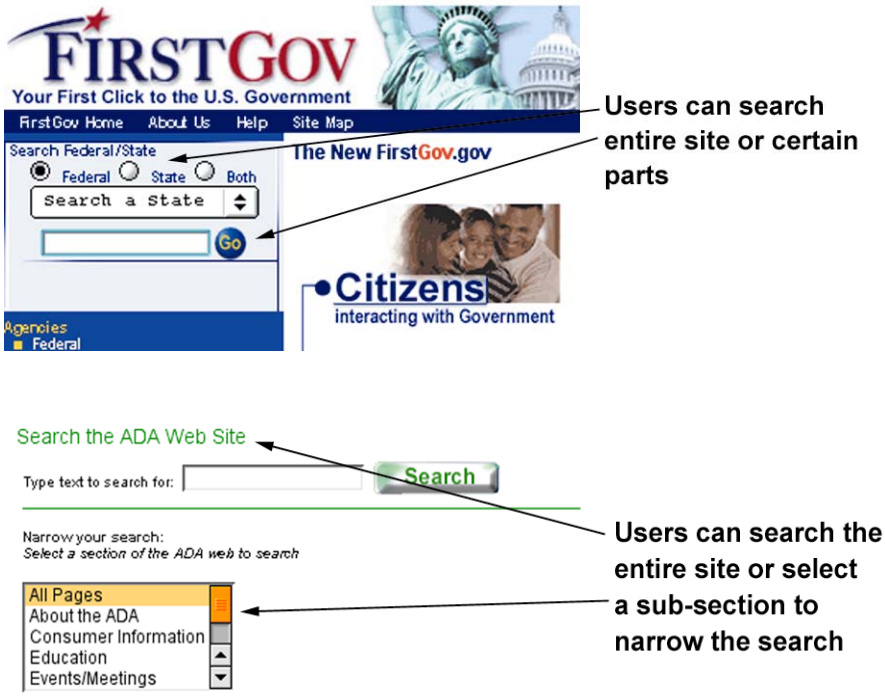

Figure 9. Examples of how users can be allowed to search a full site or a subsite to help narrow the search.

♦ Advanced Search Features Do Not Help Users

A study of Excite users (Spink, et.al., 1999) was conducted to better understand their information searching behaviors. Excite is a search engine (like Yahoo, Google, etc.) that is used to search for information throughout the Web. A total of 316 Excite users responded to the interactive survey. They provided information on a variety of search-related topics. The most interesting finding was that the results showed clearly that users tended to employ very simple search strategies, usually by entering one or two words into the search box. They did not take advantage of the advanced search capabilities, such as Boolean operators, query modifiers, etc.

One of the more advanced ways to search is to use Boolean operators, which provide ways to perform more analytical search strategies by allowing the use of "and," "or," "not," etc. Wildemuth and Downs (1998) attempted to determine whether a Boolean search would produce better results than a traditional search.

Different approaches may be appropriate for different purposes and different audiences. Traditional search methods (using one or more keywords) may be best for finding general information, such as reviewing the titles of news stories, finding relevant information in an encyclopedia, etc. Boolean search may work best when trying to find a specific answer or when searching an online library catalog.

They compared the two modes by having medical students access a biomedical database to answer eight clinical problems. In this task they found no reliable difference between the two methods. However, the participants preferred using the Boolean method.

Nielsen (2001) found in his usability studies that typical users do not have, or elect not to use, advanced search skills. Most users do not seem to be able to use advanced search or Boolean query syntax. Nielsen reported in a recent search study that the average query length was only about two words. Other studies also showed a preponderance of simple searches. Designers should ensure that their search facility could effectively handle one-word and two-word queries.

Spool (2002) has pointed out that even when advanced features are available, users rarely used them. Even when advanced features are used, they do not seem to improve success.

Advanced search capabilities, if made available at all, should be clearly defined as such.

♦ Search Results Should be Clearly Presented and Organized

Many users can have trouble interpreting the results of a search (Spink, Bateman and Jansen, 1999). The results should be clearly presented with the appropriate amount of detail and organized so that they can be readily understood.

Users almost never look beyond the second or third page of search results. Designers should prioritize search results in a useful way so that the most important hits appear high on the first page.

♦ Search Engine Boxes Should Be Designed to Accommodate Typical Queries

Search boxes should be a type-in field and not a link, and the entry field should be wide enough to accommodate a typical query. If the box is too small, the query will have to scroll horizontally, and users will not be able to see and verify all entered information.

SUMMARY AND CONCLUSIONS

This report attempts to summarize the available research on how users access information within Web sites. As more research becomes available on this topic, new information will be added to this report.

Below is a summary of the key points in the research. They are written to help frame possible Web design guidelines as they relate to searching and linking.

- **Users Have No Predisposition to Searching or Linking Behaviors -** Users do not seem to be predisposed to using either linking or searching behaviors; the behavior used appears to depend on the site's content and associated tasks. Therefore, include both links and a search capability on all key pages of a Web site.

- **Designers Need to Accommodate Both 'Linking' and 'Searching' Behaviors -** When users are searching for information or a product that is specific and unique, searching is used more frequently and tends to be more successful than linking. Therefore, for specific, unique items, like a specific book or CD title, encourage users to use the search capability. This might be accomplished by how prominently you position the search capability throughout the site, and/or how you describe search strategies in sections like 'How to Use This Site.'

  When users are searching for information or a product that is *not* specific and unique, linking is used more frequently and tends to be more successful than searching. Therefore, for non-specific, non-unique items, such as a report on thyroid cancer, users should be encouraged to use links. This might be accomplished by how prominently designers position the search capability throughout the site, and/or how they describe search strategies in sections like 'How to Use This Site.'

- **Users Are More Effective 'Linking' to Content -** Linking requires less effort from users, and tends to get them to desired content faster. To encourage linking versus searching to find products or information, provide more links on pages. Links should be descriptive, distinguishable from related links, and clearly represent users' needs.

- **Designers Should Plan for Linking Instead of Searching to Ensure Users Find Target Information, View More Related Information -** By using linking instead of searching to find information, users view significantly more related site content. Therefore, to encourage users to see and read (deal with) related information, encourage users to use linking, not searching.

- **Using Natural Language Search Capabilities Is More Efficient-** Natural language search capabilities allow people to ask questions using plain English.

These search capabilities use the information available on a company's Web site to answer questions, and have the ability to help narrow and focus the search by asking questions.

- **More Effective Search Capabilities Can Be Accomplished by Following These Guidelines:**

  - Designers should make every effort to ensure that users get relevant results on their first search attempt. Search success goes down after each attempt.
  - Designers need to be aware of, and make provision for, the terms (keywords) that users typically will use for searching.
  - Identify commonly used search terms by using surveys, search engine logs, and other techniques, and incorporate the search terms in the search engine algorithm to improve hit rates.
  - Designers should provide search templates (collections of pre-selected keywords) to help users better formulate their search queries and improve hits.
  - To help ensure search success, search engines should accommodate the following errors: misspellings, inappropriate case, spaces and punctuation, misused plurals, and typing errors.
  - Do not rely on advanced search features to improve the effectiveness of a search capability.
  - Prioritize search results in a useful way so that the most important hits appear high on the first page.
  - To help limit the scope of searches on large Web sites, provide search capabilities on certain subsites.

REFERENCES

Blinn, B. and Biers, D.W. (1999), Searching hard-copy (paper) vs. electronic documents: Role of experience, amount of text displayed, and book metaphor, *Proceedings of the Human Factors and Ergonomics Society - 1999*, 447-451.

Fang, X. and Salvendy, G. (1999), Templates for search queries: A user-centered feature for improving web search tools, *International Journal of Human-Computer Interaction,* 11(4), 301-315.

Furnas, G., Landauer, T., Gomez, L. and Dumais, S. (1984), Statistical semantics: Analysis of the potential performance of keyword information systems, In *Human Factors in Computer Systems*, Edited by John Thomas and Michael Schneider, 187-242.

JeevesOne-2002 - http://www.etestinglabs.com/main/reports/jeevesone.pdf

Nielsen, J. (2001), Search: Visible and simple, *Alertbox - useit.com*, May 13.

Nielsen, J. (1997), Search and you may find, *Alertbox - useit.com*, July 15.

Ojakaar, E. and Spool, J.M. (2001), Getting them to what they want, *User Interface Engineering Report*.

Spink, A., Bateman, J. and Jansen, B.J. (1999), Searching the Web: A survey of Excite users, *Internet Research: Electronic Networking Applications and Policy*, 9(2), 117-128.

Spool, J. (2002), In search of the perfect search: Building the perfect on-site search, *CHI 2002 Tutorial*.

UIE (2001a), Searching vs. linking, *User Interface Engineering Newsletter*, world.std.com/~uieweb/articles/, December 4.

UIE (2001b), Users don't learn to search better, *User Interface Engineering Newsletter*, world.std.com/~uieweb/articles/, November 27.

UIE (2001c), People search once, maybe twice, *User Interface Engineering Newsletter*, world.std.com/~uieweb/articles/, November 20.

UIE (2001d), Searching behaviors, *User Interface Engineering Newsletter*, world.std.com/~uieweb/articles/, May 13.

Wildemuth, B.M., Friedman, C.P. and Downs, S.M. (1998), Hypertext versus Boolean access to biomedical information: A comparison of effectiveness, efficiency, and user preferences, *ACM Transactions on Computer Human Interaction*, 5(2) 156-183.